Strategy
for
Testing, verifying and validating sounder data at JPL

Bjorn Lambrigtsen, Qing Yue

Jet Propulsion Laboratory, California Institute of Technology

AIRS Science Team Meeting
Caltech, April 25, 2018

Motivation

We aim to be a clearinghouse for sounder retrieval systems & data

- Unbiased assessments
- Support AIRS, SNPP SIPS, others
- Assist algorithm developers
- Assist software developers
- Advise sponsors

Broad expertise built on years of experience with AIRS

- Testing new algorithms
- Facilitating software implementation
- Supporting data production
- Supporting users
- Validation

Tools and processes

- Well exercised analysis processes
- Significant holdings of reference data

Document describing all this will be released shortly

Some definitions

Testing

– Is there anything wrong with it?

Verification

– Does it meet requirements?

Characterization

– How does it perform?

Validation

– How does it perform relative to “truth”?

Questions to be answered

Is it ready to be implemented?

- Purpose: Identify problems, shortcomings & issues; sanity check; iterate development
- Customer: Algorithm and software developers
- Analysis: Low to moderate complexity
- Deliverables: Internal presentations (viewgraphs)

Is it ready to be delivered/distributed/published?

- Purpose: Determine if it meets expectations/requirements; basic performance
- Customer: Algorithm developers, Project, implementers, users
- Analysis: Medium complexity
- Deliverables: Interim report; release guide

Is it better than “X”?

- Purpose: Compare retrieval systems; compare with other systems
- Customer: Algorithm developers, sponsors
- Analysis: Simple to complex
- Deliverables: Internal report

How good is it?

- Purpose: Characterize performance against a variety of comparative data
- Customer: Users
- Analysis: Moderate to complex
- Deliverables: External reports; peer reviewed papers

Levels of complexity of analysis

Stage-0

Goal:
Quick general data quality examination on key variables.

Approach
- Comparison between versions/systems
- Comparison with reanalysis

Examples
- Retrieval yield in AIRS V7 and CrIMS products
- L2 T and Q profile differences with ECMWF
- L3 TPW comparison with AMSR
- L3 Surface skin temperature and 2m temperature comparison with ECMWF

Stage-1

Goal
- Validation of key product retrievals
- Identifying possible causes of reduced retrieval performance.

Approach
- Comparison with well-developed reference datasets and tools at JPL
- Pixel-scale collocation
- Cross-relationships of multiple variables

Examples:
- Relationship between retrieval yield and surface condition/cloud condition.
- Comparison with radiosonde measurements on T and Q over land (IGRA) and over ocean (MAGIC).
- L2 near surface T and Q comparison with mesonet (over land) and shiptrack/buoy measurements (over ocean)

Stage-2

Goal
- Validation and quality check of a wider range of variables as requested by users

Approach
- Same with Stage-0 and Stage-1
- Requesting additional reference data collection and tool development.

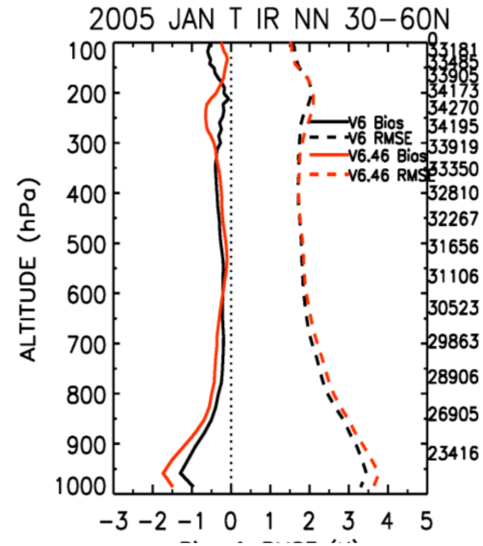
Examples
- Trace gas products: CO_2, O_3, etc
- Retrieval information content analysis: AK, DOF, vertical reso.
- Longterm trend and climate extreme.
- Comparison with new field campaign measurements: HS3 and SOCRATES

“Stage 0” level of complexity

This stage aims at examining the \general quality of the retrieval data products. They often involve only a few key parameters, and is used to determine if a new version of the retrieval system is improved or fixes a known problem. Therefore, Stage 0 often involves comparisons between versions of the system(s), and comparisons between mean states from the retrieval and from the reanalysis. These tests can typically be done quickly and repeatedly.

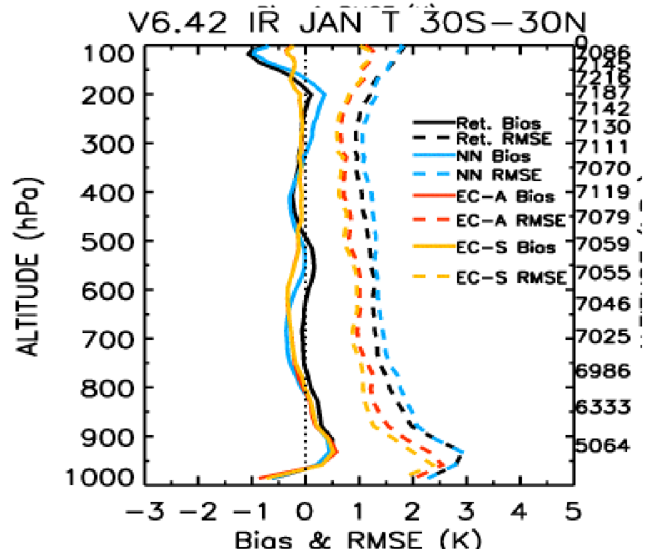
Stage 0a

- Comparison with previous version; between systems operating on same data
- Examples: AIRS V6.28 vs. AIRS V6.42; CHART(CrIMSS) vs. CLIMCAPS (CrIMSS)
- Matches are guaranteed, only QC matchup processing required



Stage 0b

- Comparison with model data
- Examples: AIRS vs. ECMWF, AIRS vs. MERRA
- Matches are guaranteed, matchup processing is relatively simple



Stage 0: Current scope

Var	Reference Data	Contributor	Method	Notes
L2 Temperature and water vapor profiles: T(P) Q(P)	Reanalysis	Qing Yue Evan Fishbein (Yue et al. 2013)	1. Collocation by the nearest neighbor method (with ability of temporal and spatial interpolation to the exact ECMWF data point). 2. Reanalysis profiles are interpolated to AIRS vertical grid. 3. Biases and RMSE are calculated for retrieval and NN from various versions of AIRS against reanalysis.	Available globally at different tim scales Initial qua check and inter-comparison
L3 Temperature and water vapor profiles: T(P) Q(P)	Reanalysis	Qing Yue Baijun Tian (Yue et al. 2013, Hearty et al. 2014)	1. Gridded data comparison on daily/monthly mean fields. 2. Yield and sampling biases. 3. Version to version changes	
L3 Total Precipitable Water Vapor (TPW)	AMSR, TMI, GMI, reanalysis	Qing Yue	1. Gridded data comparison on daily/monthly mean fields. 2. Yield and sampling biases. 3. Version to version changes.	No land da from these microwave instruments
L3 Surface Skin Temperature and Surface Air Temperature	Reanalysis	Qing Yue Evan Manning	1. Gridded data comparison on daily/monthly mean fields. 2. Yield and sampling biases. 3. Version to version changes.	Performan on these retrievals lin to the surfa type and surface emissivity, especially over frozen surfaces.

“Stage 1” level of complexity

Stage 1 testing involves comparisons with well-validated reference datasets such as in-situ measurements (ground-based and airborne) and observations from other satellites that offer higher quality retrievals of the target variables. Moreover, the computational infrastructures are either currently well-developed (Stage 1-A) or require limited efforts to update (Stage 1-B). On the other hand, the results of the comparisons with reanalysis data are further diagnosed to identify potential causes of changes among different retrieval algorithms. This stage of activities tests a more comprehensive set of parameters over a wider range of space and time to determine if resources can be devoted to implement a new version of the retrieval system and produce substantial amounts of testable data.

Stage 1a

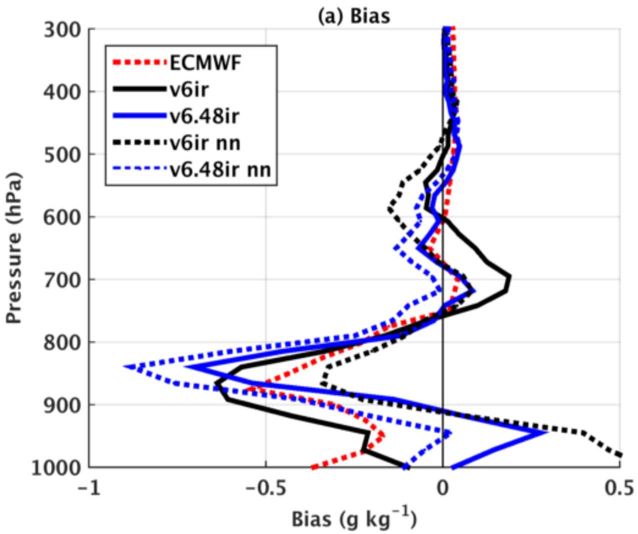
- Comparison with correlative data: limited scope, limited parameters
- Examples: T(z), q(z); limited areas with operational raobs matches only (e.g., Europe)

Stage 1b

- Comparison with correlative data: wide scope, complete set of parameters
- Examples: all variables; globally representative sampling

Stage 1c (not yet implemented)

- Comparison with correlative data: select regimes, select parameters
- Examples: ozone hole, Antarctica; T, q, tropical cyclones



Stage 1a: Current scope

Var	Reference Data	Contributor	Method	Notes
L2 Temperature and water vapor profiles: T(P) Q(P)	Dedicated sonde, IGRA sonde, and ECMWF mainly over Europe, and N. America	**Sun Wong (Wong et al. 2015)**	1. Collocate AIRS with reference data using the nearest neighbor method with temporal tolerance of 3 hours and spatial tolerance of 200 km. 2. Radiosonde data are interpolated to AIRS 100 pressure levels. 3. Biases and RMSE are calculated for retrieval and NN from AIRS and ECMWF against sonde. 4. Results are stratified by cloud fraction, surface condition, and latitude bands.	Stage 1- Sonde density s Fig. 1. Long-ter availabili reference data
	MAGIC (9/2012–10/2013) sonde and ECMWF over Pacific subtropical ocean	**Peter Kalmus Evan Manning (Kalmus et al. 2015)**	1. Collocation by the nearest neighbor method with temporal tolerance of 6 hours and spatial tolerance of 200 km. 2. Radiosonde data are interpolated to AIRS 100 pressure levels. 3. Biases and RMSE are calculated for retrieval and NN from AIRS and ECMWF against sonde. 4. Results are stratified by longitude bins.	Stage 1- Subtropi low cloud region in Northeas Pacific or (ship trac shown in 2). 9/2012–10/2013
	Reanalysis	**Qing Yue Evan Fishbein (Yue et al. 2013)**	1. Collocation by the nearest neighbor method (with ability of temporal and spatial interpolation to the exact ECMWF data point). 2. Spatial distribution of differences between AIRS (retrieval and NN) and reanalysis. 3. Sorting differences by multiple conditions: cloud, surface, season, etc to diagnose the cause of changes.	Stage 1-

Stage 1b: Current scope

L2 Near surface air temperature and water vapor: NSAT and NSWV	• Ocean: ICOADS, Buoy and shiptrack data • Land: mesonet over land (CONUS)	**R. Chris Wilson**	1. Collocation by the nearest neighbor method with temporal tolerance of 1 hour and spatial tolerance of 50 km. 2. Biases and RMSE are calculated for retrieval and NN from AIRS against reference data. 3. Results conditioned on cloud and regions.	□Stage 1- □MesoNet over CONU (Fig. 3a). □ICOADS distribution over ocean Fig. 3b). □Long-ter availability reference data.
Ozone: O_3 (Total and profile)	O_3 measured by uplooking UV-Visible spectrometer from Dumont d'Urville station (Fig. 4)	**Evan Fishbein**	1. Two closest matches in AIRS data by the nearest neighbor method. 2. Location is selected for its largest variability along the edge of the hole, near the Antarctic coast, showing the influence of stratospheric weather on polar vortex isolation and mixing 3. Specific year is identified when O_3 at this location is different from the mean climatology. 4. Time series of total O_3 and individual vertical profiles are examined.	□ Stage 1- □ Long-ter availability reference datasets to increase sample siz and site numbers.
L2 Total precipitable water: TPW	GPS ground stations	**Qing Yue Evan Manning Evan Fishbein**	1. Collocation by the nearest neighbor method with temporal tolerance of 0.5 hour and spatial tolerance of 100 km; multi-year 2. Collocation by the box method. 3. Biases and RMSE are calculated for retrieval and NN from AIRS against GPS. 4. Results conditioned on cloud, land-only	□ Stage 1- □ Possibilit to include more GPS network to cover glob land region Currently results are over US or (Fig. 5).

“Stage 2” level of complexity

This level of testing covers variables not included in Stage 0 and 1 such as surface classes, cloud-cleared radiances, cloud properties, and other trace gases. This type of testing may require further development of software and acquisition of additional reference data. Analyses that aim at information content, climate variability, long-term trend, and extreme conditions are also covered in this level.

Stage 2: Current scope (L2 products)

Var	Reference Data	Contributor	Method	Notes
L2 Cloud-cleared radiances: CC-Rads	MODIS clear radiances	R. Chris Wilson Mathias Schreier (Schreier et al, 2018)	1. Collocate multiple MODIS pixels within one AIRS FOV and only the clear MODIS pixels as flagged by the MODIS clear 35 flag are used in the analysis. 2. The AIRS CCRs are spectrally convolved to MODIS channels 22, 24, 28, 32, 33, 34, and 35, while clear MODIS radiances are spatially convolved to the AIRS field of regard. 3. Brightness temperature differences between AIRS and MODIS are calculated and compared with the expected errors indicated by the QCs of AIRS CCR product.	☐ Require collocation between sounder ai cloud imag
L2 Temperature profiles and bias drift with time	PREPQC radiosonde	Fredrick Irion	1. Collocation by the nearest neighbor method with temporal tolerance of 1 hour and spatial tolerance of 100 km. 2. Both direct comparisons based on linear interpolation and the kernel smooth method are used. 3. Both the temperature bias/RMSE and trend of T bias profiles are calculated. 4. Results are stratified by latitude.	☐ Long-ter availability reference data. ☐ Results a dominated large samples o Europe (Fi 6). ☐ Drift of bi tested.
L2 information content and vertical resolution analysis	None	Evan Fishbein Fredrick Irion	1. Information content analysis: averaging kernel, degree of freedom, retrieval error estimation 2. Vertical resolution and sensitivity	☐ Results c AIRS V6 Ozone are available a can extenc other profil retrievals.

Stage 2: Current scope (L3 products)

L3 Tropopause height, pressure, and temperature	GPS RO	**Baijun Tian**	1. Mean field and yield analysis 2. Version to version changes	
Cloud Properties	None	**Brian Kahn**	1. Pixel-scale comparisons on cloud properties including thermodynamic phase, cloud fraction, cloud top pressure among different versions.	
L3 Total Column Ozone (Daytime)	OMI	**Fredrick Irion**	1. Mean field and yield analysis 2. Version to version changes	
Other trace gases	unknown	**Vivienne Payne**	Collaboration with the composition group	

Additional reference data

Typically used to support Stage 1c testing
More will be added as needed

Campaign	Location	Time	Climate Regimes
HS3 (Hurricane and Severe Storm Sentinel)	Lat: 10 ~ 50 Lon: -160 ~ -19	Aug and Sep from 2011 to 2014	Midlat and Tropic oce severe storm
SHOUT (Sensing Hazards with Operational Unmanned Technology)	Lat: 10 ~ 50 Lon: -160 ~ -19	Aug-Sep, 2015 Feb, 2016 Aug-Oct, 2016	Midlat and Tropic oce severe storm
WISPAR (the Winter Storms and Pacific Atmospheric Rivers)	Lat: 0 ~ 90 Lon: -170 ~ -120	Feb-March, 2011	Atmospheric Rivers, Ar environment
VOCALS (VAMOS Ocean-Cloud-Atmosphere-Land Study)	Lat: -30 ~ -15 Lon: -90 ~ -70	Oct-Nov, 2008	Southeastern Pacific cloud region
SOCRATES (Southern Ocean Clouds Radiation Aerosol Transport Experimental Study)	Lat: -70 ~ -30 Lon: 130 ~ 180	Jan-Feb, 2018	Southern Ocean